基于 GEP 的高速公路通行费预测方法研究 *

刘　宁，黄樟灿，谈　庆

(武汉理工大学 理学院, 武汉 430070)

摘　要：高速公路通行费未来收入状况的预测对于高速公路运营管理、建设规划有着重要的指导意义。然而，通行费收入水平的变化受到多方面因素的影响，具有较强的非线性和复杂性，传统预测模型无法准确表达通行费收入的发展规律。针对复杂的高速公路通行费预测问题，建立了基于基因表达式编程算法(GEP)的高速公路通行费预测模型。该模型利用 GEP 算法建立通行费当前收入与历史数据之间复杂的函数关系，准确地刻画通行费收入随时间的发展规律。此外，针对节假日期间通行费减免政策的影响，提出了有效的修正模型。最后，采集了浙江沪杭甬高速公路股份有限公司等 12 家公司通行费收入的历史数据进行仿真实验，对比传统的 ARIMA 以及神经网络预测模型，结果充分验证了本文算法的有效性和准确性。

关键词：通行费预测；基因表达式编程；非线性；函数优化

中图分类号：TP391.9　　doi: 10.3969/j.issn.1001-3695.2018.01.0023

Research on freeway toll prediction method based on GEP

Liu Ning, Huang Zhangcan, Tan Qing

(*School of Science*, *Wuhan University of Technology*, *Wuhan* 430070, *China*)

Abstract: The prediction of the future income of highway toll has great guiding significance for the management and construction planning. However, the change of toll income is influenced by many factors. It has strong nonlinearity and complexity. The traditional prediction model cannot accurately express the development law of the toll income. In this paper, a highway toll prediction model based on gene expression programming algorithm (GEP) is established. The GEP algorithm is used to establish a complex functional relationship between current income and historical data, which accurately characterize the development rule of toll income over time. In addition, an effective correction model is proposed for the influence of toll reduction policies during holidays. Finally, this paper collects the historical data on the toll revenue of 12 companies such as shanghai-hangzhou-ningbo Expressway Co. , Ltd. Compared with traditional ARIMA and neural network prediction model, and the results fully verify the effectiveness and accuracy of the proposed algorithm.

Key words: toll forecast; Gene Expression Programming; nonlinear feature; function optimization

0　引言

时间序列预测方法是一类典型的预测问题，在实际生产生活中有着广泛的应用。本文基于高速公路通行费收入的时间序列样本数据，对通行费收入变化的预测问题展开研究，对指导高速公路投资者和相关职能部门进行运营管理、发展规划以及局部建设有着重要的意义[1]。

高速公路通行费由通过收费站点的车辆产生，由通行车辆的类型和数量所决定。同时，高速公路通行费收入水平的变化情况复杂，与气候、天气、经济发展、环境和政策等多方面的因素有关，传统的时间序列预测模型在该预测问题上表现较差[2]，建立有效的预测模型提高预测精度是一项具有挑战性的难题。在众多的影响因素中，收费政策的影响具有突变性和不可预测性，给通行费收入水平的预测带来了困难。2012 年起，根据有关部门的相关政策，在 4 个重大节假日(春节、清明节、劳动节、国庆节)期间，7 座以下(含)小型客车免收通行费，从而造成对应时间段内通行费收入变化趋势发生不规律的剧烈波动，这也是造成最终预测精度产生偏差的重要因素之一。

在已有的高速通行费预测研究中，主要是基于对高速公路车流量的预测来间接实现通行费预测。对于交通流车流量的预测问题，经典预测模型包括有：历史平均法[3]、指数平滑法[4]和卡尔曼滤波[5]等。然而，上述方法原理比较简单，都是基于样本数据的统计特性出发进行预测建模，预测精度较差。同时这类方法还有一个假设前提，即未来的数据变化与过去的数据具有

收稿日期：2018-01-16；**修回日期：**2018-03-16　　**基金项目：**国家自然科学基金资助项目（61672391）

作者简介：刘宁（1994-），女，浙江嵊州人，硕士，主要研究方向为智能计算；黄樟灿（1960-），男（通信作者），浙江嵊州人，教授，博导，主要研究方向为智能计算、图像处理（huangzc@whut.edu.cn）；谈庆（1991-），男，湖南长沙人，硕士，主要研究方向为图像处理.

相同的特性。但是，在高速公路预测问题中，由于节假日减免政策所造成的通行费收入突变是无法通过上述模型的简单原理来刻画的。

为了提高通行费预测精度，人们提出了更加复杂的非线性模型和智能预测模型，包括非参数回归预测方法[6]和神经网络预测方法[7]等。文献[8]中提出了一种基于神经网络的改进的自回归积分滑动平均模型(ARIMA)预测方法，该方法将预测分为两个部分，首先通过 ARIMA 模型对模型进行预测，其次利用神经网络预测方法对残差部分进行预测，通过结合两种算法来提升预测精度。但是这种方法计算相对复杂，并且存在容易陷入局部最优、计算量大的缺点。文献[9]中针对线性预测方法在交通流复杂变化趋势预测问题上的局限性，提出了一种改进的卡尔曼滤波算法，基于神经网络算法增强算法对非线性数据的预测能力。文献[10]则设计了一种基于支撑向量机(SVM)方法的交通流预测模型。上述通过交通流预测模型的通行费预测方法不易获取历史数据，同时通行费与通行车辆类型、费率等有关[11]，间接预测会产生较大误差。对此，文献[12]中提出了一种组合预测模型，结合了多种预测模型，直接对通行费收入数据进行建模和预测。

同时，现有研究主要基于一般情况下的高速公路交通流量预测结果进行交通费预测，很少考虑到重大节假日的影响。文献[13]中，作者在灰色马尔可夫预测模型的基础上提出了一种修正算法，通过计算预测残差的概率和状态来提升预测精度。文献[14]针对节假日高速公路的交通流的特性，提出了 EMD 和 GS-SVM 融合的交通流量预测模型，通过对比分析验证了融合算法的准确性与适应性。但是，以上研究方法依旧是基于交通流量的预测，通行费预测结果仍存在较大误差。

基因表达式编程算法(GEP)在复杂函数建模方面表现出较高的拟合精度和较快的收敛速度，使其具有更强的解决问题能力。GEP 建模方法能够根据时间序列数据的特征，挖掘发现数据之间的函数关系，并建立相关预测模型。文献[15]基于 GEP 算法提出了人口预测模型，该模型可以在样本少的情况下挖掘人口发展的复杂非线性模式，有效防止了过度拟合现象的发生，提供了更为准确、合理的拟合及预测结果。文献[16]针对股票对象的特点，提出了适应股票规律的 GEP-STOCK 模型，对股票涨跌情况作出了预测。针对高速公路通行费预测的复杂性，本文提出了一种基于 GEP 算法的混合预测模型来刻画通行费收入的变化规律，同时针对节假日限免政策对通行费收入的影响建立修正模型，提升算法的预测精度。

1　基因表达式编程算法

基因表达式编程算法(GEP)在 2001 年由 Ferreira 首次提出[17]，是一种新的基于基因型和表现型的自适应演化算法。GEP 表现出其先天的自组织、自适应和自学习的特性，在复杂函数的自动建模能力上具有更强的效果[18]。本文利用 GEP 算法对历史数据建立当前待测时刻通行费的映射函数，根据通行费数据变化的趋势性增长求解下一时刻的通行费预测值。

1.1　基因结构和编码方式

GEP 中的个体（或染色体）由一个或若干个等长的基因序列构成，基因编码由头尾两部分组成，运算符和终结符构成头基因，其长度 h 根据实际需求而确定，尾基因只包含终止集，尾部长度 t 的计算表达式如公式(1)所示：

$$t = h*(n-1)+1 \tag{1}$$

其中，n 表示运算符集中需要的变量个数。定义函数集 $F=\{Q,*,/,-,+\}$，终止集为 $T=\{a,b\}$，给定 $h=15, n=2$，则根据函数集 F 和终止集 T，得到基因编码“*b+a-aQab+//+b+**babbabbbababbaaa**”，如图 1 所示：

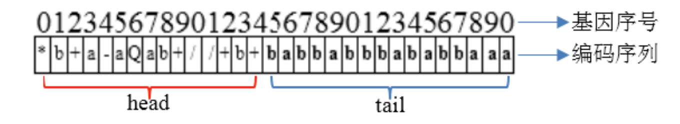


图 1　基因编码示意图

图 1 中的基因编码对应表达式树结构如图 2 所示，基因终止于第一段的 7 号位点。

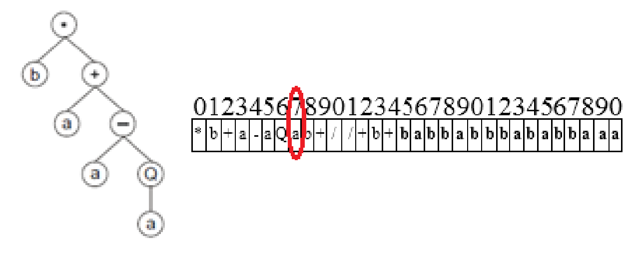


图 2　基因树结构图

1.2　GEP 遗传算子

为了使种群往更好的方向发展，在产生新个体和接受新结构的过程中，GEP 需要借助遗传算子来优化种群。GEP 算法有多种在基因序列上进行操作的遗传算子[18]，主要包括变异、插串和重组。

1)变异

变异操作通过变异概率选择个体，变异的位置也是通过随机选择，变异点所处的位置不同，变异操作也不相同，同时会对个体的表达结构产生不同的结果。

2)基因插串

插串操作对应于遗传算法中的变异，通过截取序列中的基因片段，插入到其他位置完成基因的遗传。需要注意的是，头部的基因片段只能插入到头部。该操作按照操作方式不同，分为 IS 和 RIS 两种方式。IS 插串操作不能选取根节点插串，RIS 插串操作智能选取根节点位置进行操作。

3)基因重组

基因重组操作的思想类似于遗传算法中基因的交叉操作。该操作的具体内容是选取两个基因中的某个片段，然后将对应位置的基因片段进行交换。基因重组按照交换节点方式有单点和两点两种不同的重组方式。

1.3　GEP 适应度函数选择

本文使用参考文献[19]中提到的基于相对误差的适应度函数，定义方式如下所示：

$$f_i=\sum_{j=1}^{C_t}\left(M-\left|\frac{C_{i,j}-T_j}{T_j}*100\right|\right) \tag{2}$$

其中：M 是根据具体情况设定的适应度值范围，C_t 为群体个数，$C_{i,j}$ 是基因个体 i 对第 j 个样本的预测结果，T_j 是样本 j 的实际目标值，所求的 f_i 即为第 i 个个体的适应度值。

1.4　GEP 的基本过程

针对实例分析需要解决的问题，确定输入和输出的组成元素以及算法的各控制参数。初始化种群并计算适应值，判断是否满足条件后进行各种遗传操作，以此来实现种群的进化。当算法满足终止条件时，输出最优个体或模型，GEP 基本过程如图 3 所示。

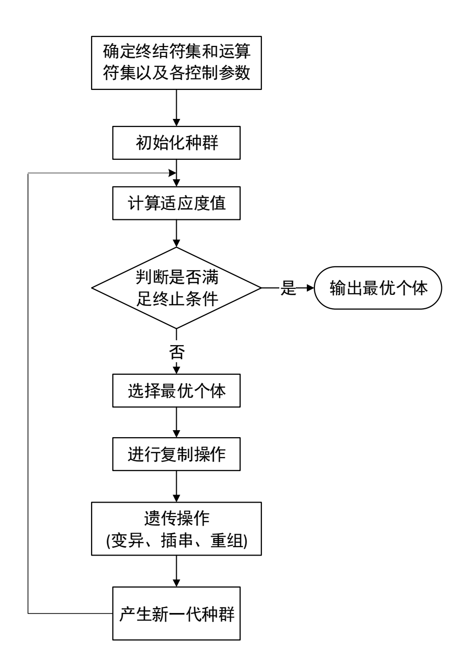


图 3　GEP 基本过程

2　基于 GEP 的高速公路通行费预测算法

2.1　高速公路通行费收入预测

本文针对通行费收入的预测问题展开研究，建立通行费当前收入数据与历史数据之间的发展关系，从而根据历史数据实现对下一时刻的通行费收入的预测。假定通行费收入的历史数据为 $\{x_0,x_1,x_2,\cdots x_n\}$，建立未来通行费与历史数据之间的函数关系，如公式(3)所示：

$$x_{n+1}=f\left(x_0,x_1,x_2,\cdots,x_n\right) \tag{3}$$

其中，如何建立准确有效的映射函数关系 $f[\bullet]$ 是本文研究的重点，该函数模型可以表现通行费收入时间序列数据的特征规律，建立优化目标函数如下式(4)所示：

$$\min\ P=\frac{1}{n}\sum_{i=1}^{n}\left|\hat{x}_i-f\left(x_{i-1},x_{i-2},x_{i-3},\cdots\right)\right| \tag{4}$$

其中 $\hat{x}_i$ 表示 i 时刻通行费的预测值。

上文中指出，通行费由通行车辆的类型和数量所决定，与气候、天气、区域文化、经济以及收费政策等多种因素相关。因此，通行费收入数据的变化具有很强的非线性和突变性，传统的预测模型无法准确描述数据的变化规律，导致最终预测结果精度不高。尤其是长期预测问题中，传统模型无法准确表达数据发展规律，预测模型不够强健，最终结果可信度较低。

针对此问题，本文在原理模型的基础上，将通行费收入金额预测问题转换为收入金额增长率的预测问题，即用增长率替代模型式(3)中的收入数据，利用 GEP 算法对增长率历史数据建立的函数关系进行通行费收入预测。

2.2　基于 GEP 的通行费收入预测模型

考虑到高速公路收费站通行费收入在节假日等其他特殊时间段具有周期性变化的特征，同时由于社会经济快速发展等因素，通行费收入序列呈现出长期趋势性，表现为在一定时期内沿着某一方向递增发展[20]。针对这一特点，本文通过差分操作，用收入增长率替代收入金额建立预测模型，其中有相邻年份对应的通行费月收入增长率数据为

$$y_i=\frac{x_i-x_{i-p}}{x_{i-p}}\ ,i=1,2,\cdots \tag{5}$$

其中：p 表示时间序列周期性变化的时间间隔(本文中 $p=12$)，根据相邻年份间的增长率存在一定的关联性和趋势性，建立预测模型如下：

$$y_{n+p}=f\left(y_{n-cp},y_{n-(c-1)p},\cdots,y_n\right) \tag{6}$$

$$x_{n+p}=\left(1+y_{n+p}\right)x_n \tag{7}$$

其中：c 表示选取的周期数(本文中 $c=2$)，y_{n+p} 表示下一时刻的收入增长率预测值，x_{n+p} 则表示对应时刻的通行费收入金额预测值，本文中采用 GEP 算法确定描述该预测过程的映射函数 $f[\bullet]$。

2.3　通行费收入预测算法

如 2.2 模型所述，通行费收入增长率的确定对预测结果起到了决定性作用，长期趋势性主要表现在过去通行费收入对当前待测时段通行费预测的影响。利用 GEP 算法对差分后的历史数据建立函数关系，选取收入增长率数据作为输入因子，收入增长率预测函数 f 作为输出因子，算法步骤如下：

输入：增长率历史数据、种群大小、基因头尾部及所有遗传算子参数。

输出：收入增长率预测模型 f 。

建模过程：

a)设置自变量、函数符集 F 和终结符集 T；

b)初始化种群，计算适应度值；

c)判断最大适应度值。若满足，转到 f)；若不满足，继续执行下一步；

d)保留最好个体，进行选择、复制、变异等遗传操作；

e)转到 b)进行循环操作；

f)满足最大适应度值或迭代次数时，输出预测模型。

计算预测值：选取历史增长率数据代入预测模型 f，并根据模型式(6)(7)得到增长率 y_{n+p} 和收入金额 x_{n+p} 的预测值。

整体预测算法框架如图 4 所示。

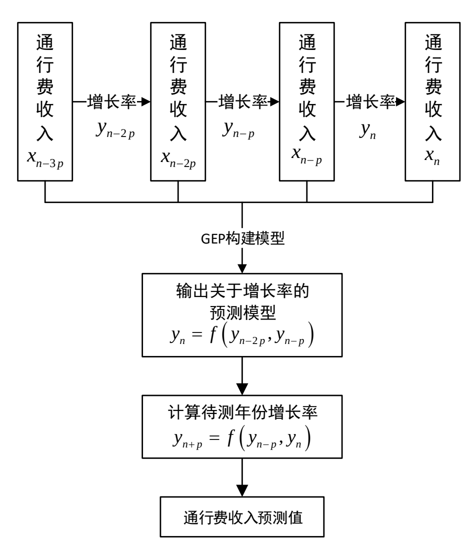


图 4 预测算法框架

2.4 通行费预测模型的局部修正

重大节假日期间的通行费限免政策会给对应月份通行费收入带来巨大的影响，特别是春节，由于其对应日期的不固定性和时间影响上的持续性，使得在对应月份的通行费收入变化趋势上的预测越发困难。针对该问题，本文在预测模型的基础上建立修正模型，对春节假期对应月份通行费收入预测进行修正。

春节假期对通行费收入的影响是大致相同的，因此以春节为基准点的通行费收入曲线的变化趋势是相同的。本文中以月收入为研究对象，因此，假期对应日期的不同对所在月份的影响也是不同的。统计 1901 到 2050 年的数据可以发现，春节对应的日期主要集中在 1 月 21 日到 2 月 20 日之间。春节时间不同，对通行费收入的影响也就不同。因此，根据春节日期与 1 月份和 2 月份间隔的长短来定义修正模型。首先选取 2 月 5 日为日期的基准点，通过计算春节日期与基准点之间的相对位置关系来对通行费收入进行修正。

因为，春节假期主要对 1、2 月份通行费收入有影响，对应的修正的通行费增长模型为：

$$\begin{cases} y_1' = 1-(D_{1,t}-D_{1,t+1})/(1+D_{1,t}) \\ y_2' = (1+D_{2,t+1})/(1+D_{2,t})-1 \end{cases} \quad (8)$$

其中：y_1', y_2' 表示修正后的通行费增长率；$D_{2,t}$ 表示第 t 年春节日期对通行费收入影响的参数。$D_{1,t}$ 的计算表达式如下：

$$\begin{cases} D_{1,t} = \xi \cdot d_{1,t} \\ D_{2,t} = \xi \cdot d_{2,t} \end{cases} \quad (9)$$

其中：$\xi \in (0,1]$ 表示春节假期对通行费月收入的影响系数。$d_{1,t}$，$d_{2,t}$ 是与春节日期相关的两个距离，定义方式如下：

a)若春节位于 1 月份，$d_{1,t}$ 表示第 t 年春节对应日期到基准点(2 月 5 日)的距离，$d_{2,t}$ 为春节日期的统计终止点(2 月 20 日)到基准点(2 月 5 日)的距离；

b)若春节位于 2 月份，$d_{1,t}$ 表示春节日期的统计起始点(1 月 21 日)到基准点(2 月 5 日)的距离，$d_{2,t}$ 第 t 年春节对应日期到基准点(2 月 5 日)的距离。

将修正后的增长率 y_1' 和 y_2' 代入 2.2 中基于 GEP 算法的通行费预测模型式(6)，得到修正的通行费预测结果。

3 仿真实验

3.1 数据来源与参数设定

针对高速公路通行费收入预测问题的研究，本文中获取了浙江省 12 家公司（包括浙江沪杭甬高速公路股份有限公司、浙江杭金衢高速公路有限公司和浙江舟山大陆连岛工程高速公路有限公司等）的月通行费收入历史数据作为研究对象，对本文预测方法的有效性进行验证。样本数据包含了 12 家公司 2012~2016 年 5 年的收入数据。如下图 5 所示，给出了上三高速公路有限公司分别在 2013、2014 和 2015 年经过标准化处理后的通行费收入数据曲线。通过对表中数据观察可以明显的发现，高速公路通行费收入曲线之间具有一定的相似性，呈现出明显的周期性和趋势性。

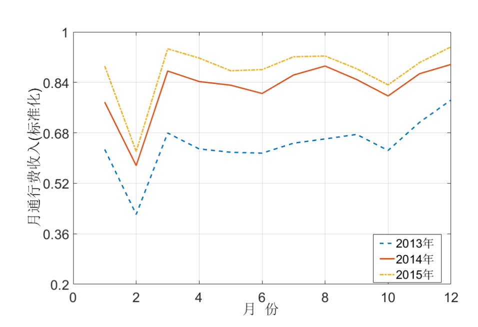


图 5 上三公司 2013 年-2015 年通行费收入数据

利用 GEP 算法对收入增长率数据进行建模预测，本文将前 48 个月份所对应的收入增长率数据作为 GEP 模拟的输入，输出未来一年 12 个月的通行费收入数据，采用式(2)作为适应度函数，参数设定如表 1 所示。

表 1 遗传参数

参数名称	参数说明
适应度函数	相对误差适应度函数
种群大小	40
最大迭代次数	10000
函数集 F	+-*/ e^x ^2
终止集 T	$\{y_{n-2p}, y_{n-p}, y_n\}$
基因个数	4
头部长度	8
连接函数	+
变异概率	0.028
基因变换概率	0.1
IS 插串概率	0.1
RIS 插串概率	0.1
基因重组概率	0.1
单点重组概率	0.3
两点重组概率	0.3

3.2　预测结果与分析

以浙江上三高速公路有限公司等 3 家公司为例，分别给出了基于 GEP 的混合预测算法所生成的模型及其对应系数值如表 2 所示。

a) 浙江上三高速公路有限公司：

$$y_n = 2(a_1 - a_2) * {y_{n-p}}^2 + e^{\left(a_3 a_4 + a_5 y_{n-2p}\right)^2} * y_{n-p} + y_{n-p} * \left(\left(1 - y_{n-2p}\right) / y_{n-2p} - \left(y_{n-2p} - y_{n-p}\right)\right) \tag{13}$$

b) 浙江舟山大陆连岛工程高速公路有限公司：

$$y_n = y_{n-p}(a_1 y_{n-2p} y_{n-p} + a_2(y_{n-2p} + 1)) + 1/y_{n-p} + y_{n-2p} - a_2 a_3 \tag{14}$$

c) 浙江金丽温高速公路有限公司：

$$y_n = \left(y_{n-p} - y_{n-2p}\right)\left(\left(a_2 + a_3\right)\left(1 - y_{n-2p}\right)^2 - a_4\right) + e^{y_{n-p} + a_1^2 y_{n-2p}^2} + a_5 y_{n-p} - a_4 + 1 \tag{15}$$

表 2　模型系数

系　数	公司简称		
	上三公司	舟山大陆连岛	金丽温公司
a_1	2.09753716	6.48548845	-7.05450911
a_2	9.31881755	-0.26039969	0.56163528
a_3	0.34159015	8.014508627	1.77278106
a_4	-6.270612335		1.88449999
a_5	8.084760236		-2.50950996

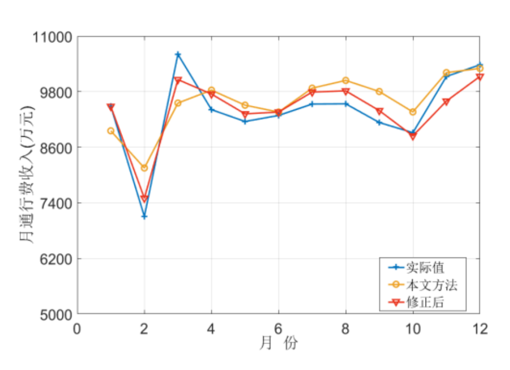
将 12 家公司的历史增长率数据 y_{n-p} 和 y_n 分别代入 GEP 对应生成的预测模型中得到增长率预测值 y_{n+p}，进一步计算得到通行费收入金额预测值，并与实际值进行比较分析。表 3 给出了 12 家公司在本文方法和传统预测方法下的未来一年通行费收入预测精度，同时给出了模型修正前后的精度对比，用年相对误差来衡量预测结果的准确程度。

表 3 不同预测方法误差比较

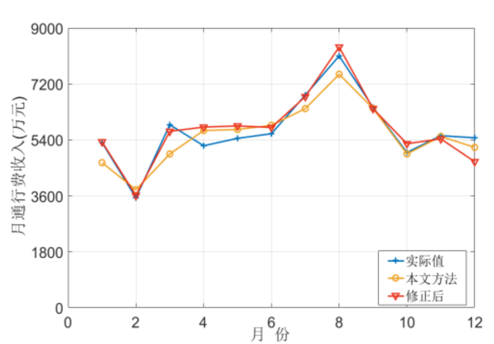
公司简称(编号)	预测方法(年相对误差%)			
	ARIMA	神经网络	本文方法	修正后
沪杭甬 JT002000001	**2.8903**	16.3557	6.583	6.0875
嘉兴 JT002002001	9.3164	10.4616	5.4047	**4.9345**
上三 JT002004001001	2.8374	3.8579	1.8258	**0.6657**
金甬 JT028000	18.5339	10.5626	2.6025	**2.1480**
申嘉湖杭(分) JT01500	3.2589	17.4046	3.1440	**2.6785**
杭金衢 JT003000	21.4959	8.8121	2.0030	**1.4979**
宁波甬台温 JT007000	13.0169	9.6133	3.3026	**3.1867**
甬台温(临海)JT005000	9.0417	15.1768	5.9624	**5.4762**
舟山连岛 JT018000	2.3962	8.4832	1.6008	**1.3849**
台州甬台温 JT006000	20.1735	6.4780	1.1714	**0.6632**
金丽温 JT008000	5.3853	5.6084	1.9020	**1.1660**
龙丽丽龙 JT010000	8.2985	8.9276	6.0908	**6.0908**

由计算结果可知，在预测方法上，除了沪杭甬公司使用 ARIMA 模型得到相对较小的误差以外，其他公司均在本文方法上有更高的预测精度。另外，在高速公路通行费收入受到强烈外部因素的影响下(例如通路、封路、旅游高峰等)，传统的预测模型 ARIMA 和神经网络表现出不稳定性。ARIMA 模型在捕捉线性关系上的局限性和对时序数据的稳定性要求是其不能作出准确预测的原因，而神经网络的不稳定性在于样本数据量不够大，神经网络的各层参数不足以保证训练的稳定。比较而言，本文算法有较强的适应性和有效性，具有相对较好的预测结果。

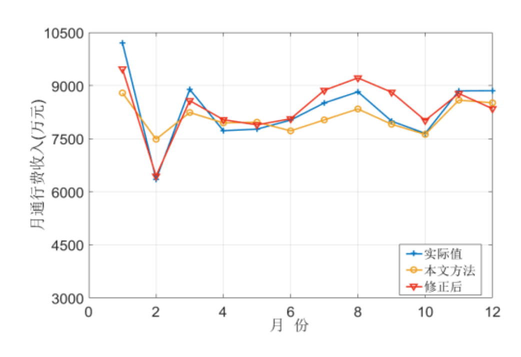
在模型修正前后的预测精度对比中发现，修正后的模型对比于原模型相对误差更小。修正模型通过分析假期出现月份的不固定性和时滞性，对当月通行费增长率进行修正，具有更高的预测精度。图 6 所示为以浙江上三高速公路有限公司等 3 家公司为例的修正前后通行费收入预测值和实际值的对比效果。



(a)浙江上三高速公路有限公司



(b)浙江舟山大陆连岛工程高速公路有限公司



(c)浙江金丽温高速公路有限公司

图 6　通行费收入预测值与实际值对比

由对比图可知，本文方法的预测值曲线和实际值曲线具有相似的趋势性和波动特征，拟合程度较好，但在春节的影响下，预测结果和实际值还存在一定差距。在模型修正后的结果中，

预测值对于真实的通行费收入在主体趋势和细节部分都拟合地更好。以图 6 中的 a 为例，可以看到在二月份时，无论是预测值还是实际值，该月份的交通收费都处于全年最低位置。说明了原有算法对该公司交通费最低值的月份预测正确，但是具体数值仍存在一定偏差。同样在 a 例中，修正算法在捕捉到最值点的同时对最值点的数值作出了更为准确的预测。

总体上，原算法能够对通行费变化趋势进行较好的预测，但在某些特殊月份上，算法过于保守并没有考虑实际环境的影响，导致关键月份的预测值存在误差。基于此本文所引入的特定月份增长率局部修正，使其在关键最值点的精确预测上有了更好的表现，进一步证明了本文基于 GEP 混合算法的有效性。

4　结束语

针对高速公路通行费预测问题，本文中基于 GEP 基因表达式编程理论提出了一种精确有效的预测方法。该方法充分利用了 GEP 算法对复杂函数关系的建模能力，确立了当前通行费收入与历史数据之间的发展关系，从而确保了预测结果的准确性。此外，针对重大节假日(以春节为例)限免政策对通行费收入水平的影响，本文在原模型的基础上进行改进，建立了有效的修正模型，从而进一步提高了预测结果的精度。对 12 家公司提供的通行费数据仿真表明，在经济发展、封路、通路和旅游高峰等各种外部因素的影响下，基于 GEP 算法建立的预测模型具有较好的适应性。最终的预测结果和误差分析表明，相较于传统的预测模型，本文方法具有更高的预测精度，可以有效的对高速公路通行费收入进行预测。

参考文献：

[1] 赵建东，王浩，刘文辉，等. 基于收费数据的高速公路站间旅行时间预测 [J]. 同济大学学报：自然科学版, 2013, 41 (12): 1849-1854. (Zhao Jiandong, Wang Hao, Liu Wenhui, *et al.* Highway travel time prediction between stations based on toll ticket data [J]. Journal of Tongji University: Natural Science, 2013, 41 (12): 1849-1854.)

[2] 潘勇. 高速公路通行费预测研究及应用 [D]. 西安：长安大学, 2015. (Pan Yong. Research and application of highway toll prediction [D]. Xi'an: Chang'an University, 2015.)

[3] Li Chisen, Chen Muchen. Identifying important variables for predicting travel time of freeway with non-recurrent congestion with neural networks [J]. Neural Computing & Applications, 2013, 23 (6): 1611-1629.

[4] 钱伟，杨矿利，杨慧慧，等. 基于组合模型的短时交通流预测 [J]. 计算机仿真, 2015, 32 (2): 175-178. (Qian Wei, Yang Kuangli, Yang Huihui, *et al.* Short-term traffic flow prediction based on combined models [J]. Computer Simulation, 2015, 32 (2): 175-178.)

[5] 余林，舒勤，柏吉琼. 基于 EMD 聚类与 ARMA 的交通流量预测方法 [J]. 公路, 2015 (5): 124-129. (Yu Lin, Shu Qin, Bai Jiqiong. Traffic flow prediction method based on EMD clustering and ARMA [J]. Highway, 2015 (5): 124-129.)

[6] Celenk M, Conley T, Graham J, *et al.* Anomaly prediction in network traffic using adaptive Wiener filtering and ARMA modeling [C]// Proc of IEEE International Conference on Systems, Man and Cybernetics. 2009: 3548-3553.

[7] Andrysiak T, Łukasz S, Choraś M, *et al.* Network traffic prediction and anomaly detection based on ARFIMA model [C]// Advances in Intelligent Systems & Computing. 2014: 545-554.

[8] 谭满春，冯荦斌，徐建闽. 基于 ARIMA 与人工神经网络组合模型的交通流预测 [J]. 中国公路学报, 2007, 20 (4): 118-121. (Tan Manchun, Feng Luobin, Xu Jianmin. Traffic flow prediction based on hybrid ARIMA and ANN model [J]. China Journal of Highway and Transport, 2007, 20 (4): 118-121.)

[9] 贺国光，李宇，马寿峰. 基于数学模型的短时交通流预测方法探讨 [J]. 系统工程理论与实践, 2000, 20 (12): 51-56. (He Guoguang, Li Yu, Ma Shoufeng. Discussion on short-term traffic flow forecasting methods based on mathematical models [J]. Systems Engineering-Theory & Practice, 2000, 20 (12): 51-56.)

[10] 徐启华，杨瑞. 支持向量机在交通流量实时预测中的应用 [J]. 公路交通科技, 2005, 22 (12): 131-134. (Xu Qihua, Yang Rui. Traffic flow prediction using support vector machine based method [J]. Journal of Highway and Transportation Research and Development, 2005, 22 (12): 131-134.)

[11] 刘闯. 高速公路运营收入预测模型与方法 [J]. 中外公路, 2005, 25 (3): 125-127. (Liu Chuang. Forecasting model and method of expressway operating income [J]. Journal of China & Foreign Highway, 2005, 25 (3): 125-127.)

[12] 赵琪. 高速公路收费站收费预测研究 [D]. 西安：长安大学, 2010. (Zhao Qi. Highway toll prediction research [D]. Xi'an: Chang'an University, 2010.)

[13] 徐蒙蒙，严凌. 基于灰色马尔可夫的节假日高速公路实施免费通行交通量预测 [J]. 物流科技, 2017, 40 (6): 101-105. (Xu Mengmeng, Yan Lin. Forecast of free traffic volume based on gray Markov's holiday highway [J]. Logistics Sci-Tech, 2017, 40 (6): 101-105.)

[14] 冯腾. 基于 EMD 和 GS-SVM 融合的节假日高速公路交通流预测研究 [D]. 西安：长安大学, 2016. (Feng Teng. Research on prediction of freeway traffic flow in holidays based on EMD and GS-SVM [D]. Xi'an: Chang'an University, 2016.)

[15] 刘萌伟，黎夏，刘涛. 基于基因表达式编程的人口预测模型 [J]. 中山大学学报：自然科学版, 2010, 49 (6): 115-120. (Liu Mengwei, Li Xia, Liu Tao. A gene expression programming algorithm for population prediction problems [J]. Acta Scientiarum Naturalium Universitatis Sunyatseni: Natural Science, 2010, 49 (6): 115-120.)

[16] 廖勇，唐常杰，元昌安，等. 基于基因表达式编程的股票指数时间序列分析 [J]. 四川大学学报：自然科学版, 2005, 42 (5): 931-936. (Liao Yong, Tang Changjie, Yuan Chang'an, *et al.* Time series analysis of stock index based on gene expression programming [J]. Journal of Sichuan University:

Natural Science Edition, 2005, 42 (5): 931-936.)

[17] Ferreira C. Gene Expression Programming: a New Adaptive Algorithm for Solving Problems [J]. Computer Science, 2001, 21 (2): 87-129.

[18] 余锡伦. 基于 GEP 的常微分方程组演化建模研究 [D]. 赣州: 江西理工大学, 2012. (Yu Xilun. Research on evolution modeling of ordinary differential equations based on GEP [D]. Ganzhou: Jiangxi University of Science and Technology, 2012.)

[19] Ferreira, Cândida. Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence. [J]. Engineering Applications of Artificial Intelligence, 2002, 1 (3): 223–225.

[20] 魏艳华, 王丙参, 郝淑双. 统计预测与决策 [M]. 成都: 西南交通大学出版社, 2014. (Wei Yanhua, Wang Bingcan, Hao Shushuang. Statistical forecasting and decision-making [M]. Chengdu: Southwest jiaotong University Press, 2014.)